

2013

OMR

*Specifikacija zahtjeva na
sistem*

IREM FILANDRA
MARKO OBRADOVIĆ



SADRŽAJ

1. UVOD	2
2. KREIRANJE FORME.....	3
2.1. PODEŠAVANJE FORME.....	4
3. PREPOZNAVANJE I VERIFIKACIJA FORME	5
4. IZLAZI SISTEMA	5
4.1. ZAHTVJEVI REZULTATA OBRADÉ.....	5
4.2. ZAHTJEVI POSTAVLJENI NA FORMULARE	6
5. NEFUNKCIONALNI ZAHTVJEVI	7

1. UVOD

OMR (Optical mark recognition) software na čijem razvoju radimo je web bazirani software koji podržava optičko prepoznavanje i automatsku obradu rezultata prepoznavanja popunjenih formi. Postojeći queXF OMR sistem se bit će platforma za razvoj našeg software-a. Ovaj software će korisnicima omogućiti :

- Dodavanje novih formi/šablona na osnovu skeniranih dokumenata u .pdf formatu i generisanje novih formi unutar sistema
- Preciznu verifikaciju popunjenih formi koja će biti potvrđena od strane operatora sistema
- Pohranjivanje i obrada podataka koji su rezultat OMR-a i OCR-a



Slika 1. Opis OMR sistema

Pored OMR-a kao primarne funkcionalnosti, software će pružati i OCR (optičko prepoznavanje znakova) čiji će izvor biti skenirane forme u .pdf formatu ili forme generisane unutar sistema.

Administrator sistema će biti zadužen za kreiranje novih formi i dodavanje novih korisnika/operatora unutar sistema.

Operator sistema će biti zadužen za verifikaciju formi koje su mu dodjeljene od strane administratora sistema.

Verifikacija će predstavljati potvrdu OMR i OCR izlaznih podataka i ako je potrebno editovanje istih podataka u slučaju da se desilo pogrešno prepoznavanje.

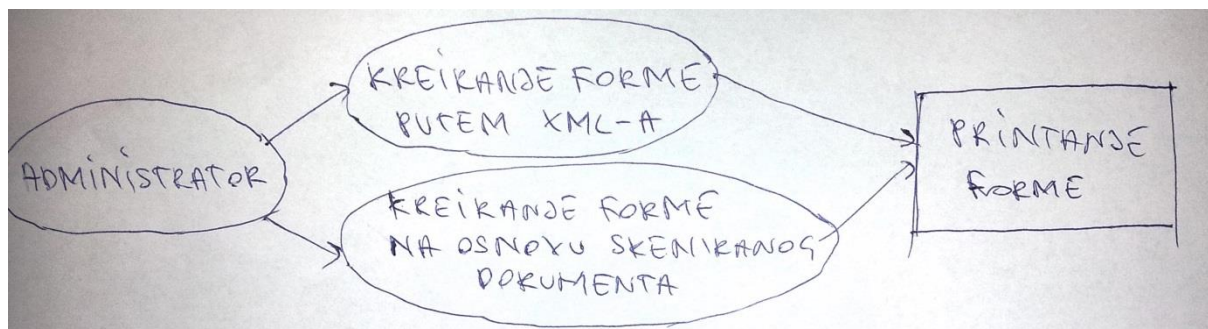
Verifikovana forma sa svim podacima će biti smještena u bazu podataka.

Naš sistem će biti baziran na klijent-server modelu gdje će klijent koristiti OMR web aplikaciju na windows ili linux platformama dok će centralni server biti baziran na linux platformi.

2. KREIRANJE FORME

Sistem će omogućiti kreiranje nove forme na dva načina (u oba slučaja ova operacija će biti izvršena od strane administratora sistema) :

1. Generisanje nove forme unutar sistema na osnovu dokumenta napisanog u XML-u
2. Kreiranje nove forme na osnovu skeniranog dokumenta koji je spašen u .pdf formatu



Slika 2. Opis kreiranja forme

U prvom slučaju sistem će na osnovu XML koda koji unosi administator unutar određene sekcije sistema generisati novu formu koja će biti spašena u .pdf formatu i pohranjena na serveru. Ako se radi o drugom slučaju sistem će na osnovu skeniranog dokumenta kreirati novu formu koja će također biti pohranjena na serveru.

Sistem će izvršavati OMR I OCR operacije nad dvije vrste formi upitnicima (tip forme gdje korisnici popunjavaju obrazac odgovarajući na određena pitanja oznakama, pisanim tekstom ili na neki drugi način) i testovima.

Ako se radi o testovima važno je da se provjerava tačnost odgovora od strane korisnika što nije slučaj kod upitnika kada je primarni cilj samo prepoznavanje i pohranjivanje rezultata upitnika. Nakon što je forma kreirana sistem će omogućiti printanje forme u određenom broju primjeraka koji su spremni za popunjavanja.

Sistem zahtjeva da svaka forma sadrži barcode (reprezentaciju podataka koji su povezani sa određenim objektom gdje je podržano optičko čitanje/prepoznavanje) tako da će barcode biti generisan od strane sistema na određenoj poziciji unutar forme.

Ako se radi u upitnicima barcode može sadržavati podatke o broju upitnika, podatke vezane za osobu koja popunjava upitnik ili neke druge podatke.

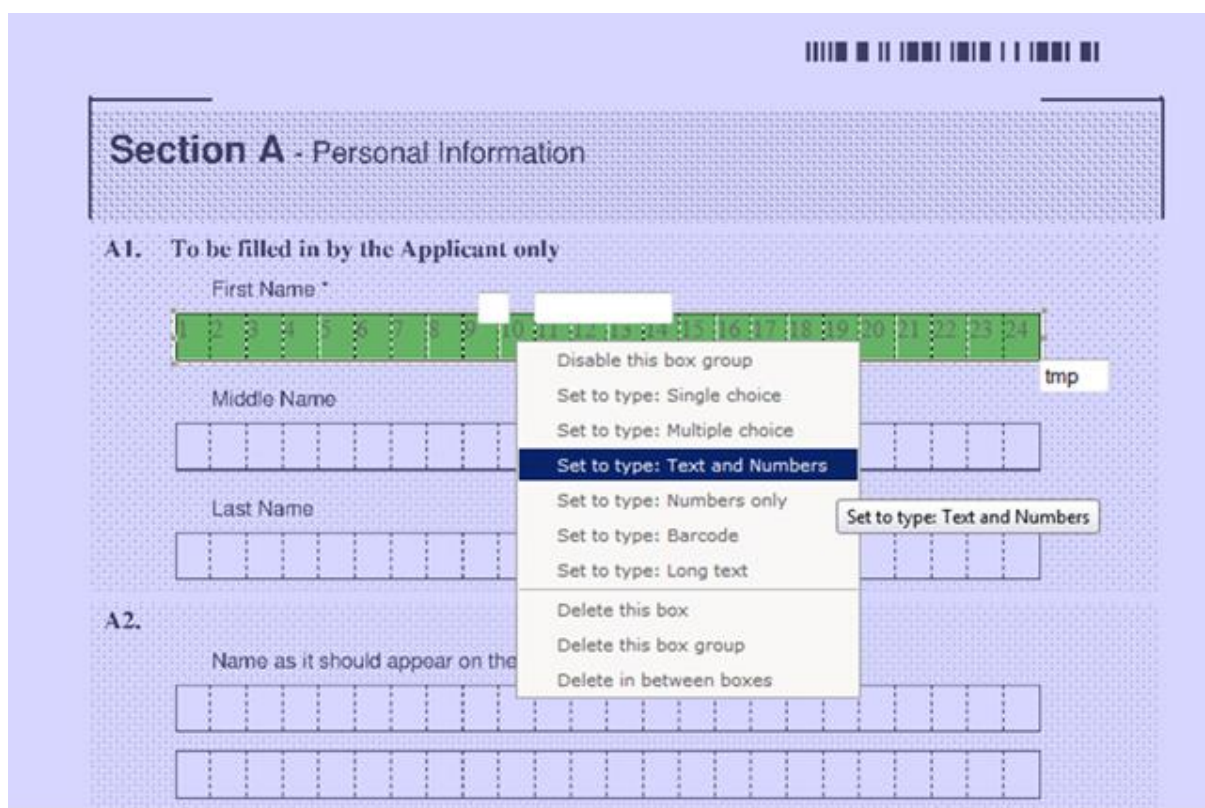
Ako se radi o testovima barcode može sadržavati podatke vezane za tačnost odgovora testova, grupaciju testova i ostalo.

Nakon što je forma kreirana zadatak administratora je podešavanje / binding forme. Korištenjem web aplikacije nad kreiranom formom administrator će mišem označiti polja nad kojim će sistem vršiti OMR i OCR prepoznavanja.

2.1. PODEŠAVANJE FORME

Nakon kreiranja, podešavanje forme se vrši označavanjem polja mišem nad kojima želimo da sistem izvrši prepoznavanje. Sistem će omogućiti imenovanje označenih blokova i postavljanje njihovih tipova na :

- tekstualne, numeričke karaktere
- oznake određene vrste npr. „x“
- barcode



Slika 3. Primjer podešavanja forme unutar queFX sistema

Sistem će omogućiti pohranjivanje forme sa svim svojim postavkama/podešavanjima tako da će pravilno podešana forma biti spremna za naknadno prepoznavanje od strane sistema.

3. PREPOZNAVANJE I VERIFIKACIJA FORME

Sistem će omogućiti prepoznavanje nad blokovima koje je kreirao administrator. Nakon što je završen proces podešavanja forme, forme su spremne za printanje i popunjavanje od strane korisnika. Popunjena forma se skenira i upload-a u .pdf formatu tako da sistem može vršiti prepoznavanje. Administrator sistema kreira korisnike operatore i dodjeljuje im određeni broj formi nad kojima će vršiti verifikaciju.

Operator pristupa sistemu sa svojim korisničkim informacijama i ima mogućnost pristupiti procesu verifikacije formi. Nakon što su izvršena OMR i OCR prepoznavanja zadatak operatora je da potvrdi i edituje (ako je potrebno) rezultate prepoznavanja.

Proces verifikacije traje kratko, sistem će omogućiti proces verifikacije koji će trajati nekoliko sekundi ako se radi o manjim formama.

Rezultat prepoznavanje i verifikacije forme su podaci koji su pohranjeni u bazu podataka sistema i sistem ih naknadno može koristiti.

4. IZLAZI SISTEMA

Izlazi ovakvog sistema se mogu grupisati u dvije kategorije – gotovi formulari ili testovi, spremni za štampanje, i rezultati obrade tih formulara. Izlazi iz sistema su svi oni dokumenti ili podaci koji napuštaju granice sistema, a koriste ih ljudi ili drugi sistemi. Imajući to u vidu, postavljaju se sljedeći sistemski zahtjevi:

4.1. ZAHTJEVI REZULTATA OBRADE

- 1.1. Sistem treba omogućiti štampanje rezultata – Nakon uspješne obrade formulara, korisniku se treba omogućiti štampanje stanja nakon obrade. Ako je u pitanju upitnik, treba omogućiti prikaz sumarnog stanja svakog predefinisiranog odgovora. U slučaju testova, sistem treba da štampa imena i broj bodova.
- 1.2. Sistem treba omogućiti pohranu rezultata u bazu podataka – Ova baza podataka omogućava trajnu pohranu dok ujedno pruža informacije drugim sistemima za daljnju obradu.
- 1.3. Baza podataka treba da pravi razliku između rezultata obrade upitnika i testova.
- 1.4. Baza podataka treba sadržavati jedinstveni kod (barcode) formulara i vrijednosti obrade vezane za njega.
- 1.5. Baza treba sadržavati listu pitanja i tačne odgovore (za slučaj testova) vezane za barcode.

4.2. ZAHTJEVI POSTAVLJENI NA FORMULARE

Formulari bi trebali sadržavati identifikaciju da bi bili pogodni za automatsku obradu podataka i njihovo upravljanje i praćenje. Moramo omogućiti korisniku da testove razvrsta u grupe - jedna grupa je jedna kombinacija pitanja npr. jedan ispit. Ovo nam daje jednu kategoriju identifikacije. Nakon toga je potrebno da unutar tog ispita identifikujemo osobu koja je radila ispit. To se postiže na način da se svakom studentu ili ciljanoj grupi predefinišu dijelovi barcode koji bi se fizički dodavali na predviđeno mjesto na formularu. Druga mogućnost je definisanje identifikacijskih polja u koja osoba koja popunjava formular upisuje ime ili za nju predefinisane cifre. Te cifre npr mogu biti broj indexa ili neki drugi već definisani jedinstveni broj. Sistem bi potom koristio OCR tehniku za prepoznavanje rukopisa.

Na osnovu ovoga, postavljaju se sljedeći zahtjevi:

- 1.6. Sistem treba generisati barcode za svaku grupu formulara.
- 1.7. Sistem treba pružiti mogućnost generisanja barcode-ova za svaku osobu definisanu unutar sistema.
- 1.8. Sistem treba omogućiti kreiranje polja za identifikaciju za OCR sistem koja se popunjavaju brojevima ili ručno pisanim tekstom.

5. NEFUNKCIONALNI ZAHTVJEVI

Nefunkcionalni zahtjevi su zahtjevi koji se postavljaju na osobine sistema, a to uključuje sljedeće zahtjeve:

1. Platformska kompatibilnost

Dijelovi sistema tj klijentske aplikacije moraju da rade na pod Windows i Linux operativnim sistemima, pa shodno tome treba razviti aplikaciju koja je cross-platform kompatibilna. Serverska aplikacija i web aplikacija će se nalaziti na Linux operativnom sistemu.

2. Arhitektura sistema

Arhitektura sistema treba biti klijent-server. Klijenti su predstavljeni klijentskom aplikacijom na kojima se obavlja unos u sistem, obrada podataka, a tako obrađeni podaci se šalju na server. Unos podrazumijeva skenirane formulare ili kreiranje formulara. Server je centralno mjesto za smještaj baze podataka i opsluživanje upita za podacima. Druga funkcija servera je i hostanje web aplikacije i opsluživanje korisničkih zahtjeva. Server je LAMP paket software-a.

3. Przistentnost podataka

Perzistentnost podataka će biti omogućena upotrebom MySQL baze podataka na serverskoj strani.

4. Sigurnost

Sigurnost podataka treba biti omogućena na način zabrane pristupa korisnicima izvan sistema, kao i razdvajanjem odgovornosti korisnika sistema. Rezultati testova mogu biti dostupni samo nadležnim korisnicima koji su zaduženi za pregled rezultata i njihovo korištenje.

5. Pouzdanost

Očekivana pouzdanost OMR i OCR alata mora biti minimalno 90%.

6. Softverski paketi

U sistemu će se koristiti open source softverski paketi za OMR i OCR kako na klijentskoj, tako i na serverskoj strani. Server je LAMP baziran – open source rješenje. Za razvoj klijent aplikacija i web aplikacije će se koristiti java i php.